

Sequence Comparison A

SEQ ID NO:23

AB005659
 LOCUS AB005659 4939 bp mRNA PRI 21-OCT-1997
 DEFINITION Homo sapiens SMRP mRNA, complete cds.
 ACCESSION AB005659
 VERSION AB005659.1 GI:2554609
 KEYWORDS SMRP.
 SOURCE Homo sapiens bone marrow cDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 4939)
 AUTHORS Suzuki,T.
 TITLE Direct Submission
 JOURNAL Submitted (10-JUL-1997) to the DDBJ/EMBL/GenBank databases.
 Toshihiro Suzuki, National Cancer Center Research Institute,
 Pharmacology Division; Tsukiji 5-1-1, Chuo-ku, Tokyo 104, Japan
 (E-mail:tssuzuki@gan2.res.ncc.go.jp, Tel:03-3542-2511,
 Fax:03-3542-1886)
 REFERENCE 2 (sites)
 AUTHORS Suzuki,T., Nishio,K., Sasaki,H., Kurokawa,H., Saito-Ohara,F.,
 Ikeuchi,T., Tanabe,S., Terada,M. and Saijo,N.
 TITLE cDNA cloning of a short type of multidrug resistance protein
 homologue, SMRP, from a human lung cancer cell line
 JOURNAL Biochem. Biophys. Res. Commun. 238 (3), 790-794 (1997)
 MEDLINE 97472289
 FEATURES Location/Qualifiers
 source 1. .4939
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="3"
 /map="3q27"
 /tissue_type="bone marrow"
 gene 737. .3577
 /gene="SMRP"
 CDS 737. .3577
 /gene="SMRP"
 /note="a short type of multidrug resistance protein
 homologue"
 /codon_start=1
 /protein_id="BAA22887.1"
 /db_xref="GI:2554610"

```

/translation="MKNATLAWDSSHSSIQNSPKLTPKMKDKRASRGKKEKVRQLQR
TEHQAVLAEQKGHLLLSDERPSPEEEEGKHIHLGHLRLQRTLHSIDLEIQEGKLVGI
CGSVGSGKTSLSIAILGQMTLLEGSSIAISGTFAYVAQQAWILNATLRDNILFGKEYDE
ERYNSVLNSCCLRPDLAILPSSLTEIGERGANLSSGGQRQRISLARALYSDRSIYILD
DPLSALDAHVGHNIFNSAIRKHLKSKTVLVFVTHQLQYLVDCCDEVIFMKEGCITERGTH
EELMNLNGDYATIFNNLLLGETPPVEINSKKETSGSQKKSQDKGPKTGSVKKEKAVKP
EEGQLVQLEEKQGQGSPWSVGVYIQAAGGPLAFLVIMALFMLNVGSTAFSTWWLSYW
IKQGSGNTTVTRGNETSVSDSMKDNPQMYYASIYALSMAVMLILKAIRGVVFVKGTL
RASSRLHDELFRRILRSPMKFDTTPTGRILNRFSKDMDEVDVRLPFQAEMFIQNVL
VFFCVGMIAGVFPWFLVAVGPLVILFSVLHIVSRVLIRELKRLDNITQSPFLSHITSS
IQGLATIHAYNKGQEFLHRYQELLDDNQAPFFLFTCAMRWLAVRLDLISIALITTTGL
MIVLMHGQIPPAYAGLAISYAVQLTGLFQFTVRLASETEARFTSVERINHYIKTLSLE
APARIKNKAPSPDWQEGEVTFENAEMRYRENLPVLKKVSFTIKPKEKIGIVGRTGS

```

BEST AVAILABLE COPY

GKSSLGMALFRLVELSGGCIKIDGVRISDIGLADLSKLSIIPQEVLFSGTVRSNLD
PFNQYTEDQIWDALERTHMKECIAQLPLKLESEVMENGDNFSVGERQLLCIARALLRH
CKILILDEATAAMDTETDLLIQETIREAFADCTMLNIAHRLHTVLGSDRIMVLAQQGV
VEFDTPSVLLSNDSSRFYAMFAAAENKVAVKG"

polyA_site 4939
/note="15 A nucleotides"
BASE COUNT 1174 a 1224 c 1249 g 1292 t
ORIGIN

Query Match 11.6%; Score 425; DB 85; Length 4939;
Best Local Similarity 56.1%; Pred. No. 2.1e-89;
Matches 800; Conservative 0; Mismatches 625; Indels 0; Gaps 0;

Qy 2217 ggtttccgctccccatgagttcttgacaccatcccaataggccggctttgaactg 2276
|| | | || | || | || | || | || | || | || | || | || | || | || |
Db 2125 GATCCTTCGAAGCCCTATGAAGTTTTGACACGACCCCCACAGGGAGGATTCTAACAG 2184

Qy 2277 ctgcgcaggggacttgaaacagactggaccagcttgcctcatcttcagagcagttcct 2336
|| | | || | || | || | || | || | || | || | || | || | || |
Db 2185 GTTTCCAAGACATGGATGAAGTTGACGTGCGGCTGCCGTTCCAGGCCGAGATGTTCAT 2244

Qy 2337 ggtcctgtccttaatggtcatcgccgtcctgttgcattgtcagtgtctgtctccatata 2396
| | || || | | | || | || | || | || | || | || | || | || |
Db 2245 CCAGAACGTTATCCTGGTGGTCTTCTGTGTGGAAATGATCGCAGGAGTCTTCCGGTT 2304

Qy 2397 cctgttaatgggagccataatcatggttatttgcatttattatatgtgttcaagaa 2456
|| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2305 CCTTGTGGCAGTGGGGCCCTTGTCATCCTCTTTCACTGCACATTGTCTCCAGGGT 2364

Qy 2457 ggccatcggtgttcaagagactggagaactatagccggctccttattctcccacat 2516
|| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2365 CCTGATTGGGAGCTGAAGCGTCTGGACAATATCACGCAGTCACCTTCCTCTCCCACAT 2424

Qy 2517 cctcaattctctgcaaggcctgagctccatccatgtctatggaaaaactgaagacttcat 2576
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2425 CACGTCCAGCATACAGGCCCTGCCACCATCCACGCCTACAATAAGGGCAGGAGTTCT 2484

Qy 2577 cagccagttaaaggactgatgcgcagaataactacacctgtctgttttatcttc 2636
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2485 GCACAGATACCAGGAGCTGCTGGATGACAACCAAGCTCCTTTTTGTTACGTGTGC 2544

Qy 2637 cacacgtggatggcattggaggctggagatcatgaccaaccttgtgaccttggctgtgc 2696
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2545 GATGCGGTGGCTGGCTGTGCGGCTGGACCTCATCAGCATGCCCTCATCACCACGGG 2604

Qy 2697 cctgttcgtggctttggcattccatccacccctactcctttaaagtcatggctgtcaa 2756
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2605 GCTGATGATCGTTCTATGCACGGGCAGATTCCCCCAGCCTATGCCGTCTGCCATCTC 2664

Qy 2757 catcgtgctgcagctggcgccagcttccaggccactgcccggattggctggagacaga 2816
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2665 TTATGCTGTCCAGTTAACGGGGCTGTTCCAGTTACGGTCAGACTGGCATCTGAGACAGA 2724

Qy 2817 ggcacagttcacggctgttagagaggatactgcagtgatcatgaagatgtgtctcgaa 2876

TEST COPY

Db	2725	AGCTCGATTACCTCGGTGGAGAGGATCAATCACTACATTAAGACTCTGTCCTTGGAAAGC	2784
Qy	2877	tcctttacacatggaaggcacaagggttgcggccagggtggccacagcatgggaaatcat	2936
Db	2785	ACCTGCCAGAATTAAGAACACAAGGCTCCCTCCCTGACTGGCCCCAGGAGGGAGAGGTGAC	2844
Qy	2937	atttcaggattatcacatgaaatacagagacaacacaccaccgtgcttcacggcatcaa	2996
Db	2845	CTTGAGAACGCAGAGATGAGGTACCGAGAAAAACCTCCCTCTTGTCCCTAAAGAAAAGTATC	2904
Qy	2997	cctgaccatccgcccacgaagtgggtggcatcggtggaaaggacggctctggaaagtc	3056
Db	2905	CTTCACGATCAAACCTAAAGAGAAAGATTGGCATTGTGGGGCGGACAGGATCAGGGAAGTC	2964
Qy	3057	ctcctgggcatggctcttccgcctggagccccatggcaggccggattctcattga	3116
Db	2965	CTCGCTGGGGATGCCCTCTCCGCTGGTGGAGTTATCTGGAGGCTGCATCAAGATTGA	3024
Qy	3117	cggcgtggacattgcagcatcgccctggaggacttgccgtcaagctctcagtgtatccc	3176
Db	3025	TGGAGTGAGAATCAGTGATATTGGCCTTGCCGACCTCCGAAGCAAACCTCTATCATTCC	3084
Qy	3177	tcaagatccagtgtgcgtctcaggaaccatcagattcaacctagatcccttgaccgtca	3236
Db	3085	TCAAGAGCCGGTGTGTCAGTGGCACTGTCAGATCAAATTGGACCCCTCAACCAGTA	3144
Qy	3237	cactgaccagcagatctggatgccttgagaggacattcctgaccaaggccatctcaaa	3296
Db	3145	CACTGAAGACCAGATTGGATGCCCTGGAGAGGACACACATGAAAGAATGTATTGCTCA	3204
Qy	3297	gttccccaaaaagctgcatacagatgtggtgaaaacgggtggaaactctctgtggggga	3356
Db	3205	GCTACCTCTGAAACTTGAATCTGAAGTGATGGAGAATGGGGATAACTCTCAGTGGGGGA	3264
Qy	3357	gaggcagctgtctgcattgccaggctgtgcattcgcaactccaagatcatccttatcga	3416
Db	3265	ACGGCAGCTCTGTGCATAGCTAGGCCCTGCTCCGCCACTGTAAGATTCTGATTTAGA	3324
Qy	3417	tgaagccacagc cttcattgcacatggagacagacaccctgatcccgccacaatccgtga	3476
Db	3325	TGAAGCCACAGCTGCCATGGACACAGAGACAGACTTATTGATTCAAGAGACCATCCGAGA	3384
Qy	3477	agcattccaggctgcaccgtgcgtcattgcccaccgtgtcaccactgtgactg	3536
Db	3385	AGCATTGCGAGACTGTACCATGCTGAACATTGCCATGCCCTGCACACGGTTCTAGGCTC	3444
Qy	3537	tgaccacatcctggatatggcaatggaaagggtggatgtcgccggaggtact	3596
Db	3445	CGATAGGATTATGGTGTGGCCCAGGGACAGGTGGAGTTGACACCCCCATCGGTCT	3504
Qy	3597	gcggaagaaggcctgggtcattgttcgcaggccctatggccacagc	3641
Db	3505	TCTGTCCAACGACAGTTCCGATTCTATGCCATGTTGCTGCTGC	3549

BEST AVAILABLE COPY

THIS PAGE BLANK (USPTO)